caBIO Wiki Home Page

Contents
- Overview of caBIO - caBIO for End Users - caBIO for Technical Users - Collaboration Opportunities - caBIO FAQ - Using the caBIO Wiki

caBIO Project Sites
- caBIO GForge Project - caBIO Python API GForge Project - caBIO AJAX with REST API Wiki - caCORE .NET GForge Project - caCORE API GForge Project - caCORE Perl API GForge Project

Quick Links
- caBIO Tools Page - caBIO Download - caBIO Home (Webservice) - caBIO Demonstration Site - caBIO (API) Users Mail List - caBIO Developers Mail List (extending the API, adding data sources, enhancing source code) - NCI Core Infrastructure Leads - Application Support - caBIO Releases - caBIO Data Sources

Overview of caBIO

Conducting biomedical research requires access to experimental data, as well as associated molecular annotations. Annotations providing detailed information on the molecular origin, biological process, and genetic alterations can provide important insight on experimental outcomes. caBIO (cancer Bioinformatics Infrastructure Objects) is a robust resource for accessing molecular annotations from curated data sources in an integrated view in support of knowledge discovery. The entities that concern the Central Dogma of Molecular Biology are the core of the model. Data for the classes come from a variety of sources, including CGAP, Unigene, the Cancer Gene Index (CGI) project, the Pathway Interaction Database (PID), as well as a number of array manufacturers (for microarray annotations). Given the dynamic nature of this information, the data in caBIO is updated semi-monthly through a series of ETL (Extract, Transform and Load) processes.

caBIO for End Users

caBIO End User Scenarios

caBIO provides the research community with access to a variety of integrated biomedical data to address molecular annotation needs. Below is an example high-level user scenarios demonstrating potential uses of caBIO.

A researcher studying the "CHEK2" gene for links to breast cancer needs a complete list of known SNPs located on this gene. caBIO can be leveraged to retrieve SNPs for a given gene symbol "CHEK2".

caBIO End User Access

The caBIO Home Page, portlet, and iPhone App provide access to caBIO data for end users.

- caBIO Home Page – The caBIO Home Page is the entry point for browsing and querying caBIO data from caBIO search criteria. The caBIO Home Page also provides a full text search capability facilitating a keyword search.
- caBIO Portlet – The caBIO Portlet is a user interface component that operates within the caGrid® Portal as a community service. The caBIO Portlet provides a simple search for performing keyword queries and a templated-based search allowing users to execute pre-defined queries. Currently, the caBIO Portlet provides support for the following templated searches:
  - *Genome Range Queries* – Absolute range queries by Taxon or Chromosome
  - *Microarray Annotations* – Reporter lookup by Gene, SNP, or Reporter Name
  - *Genomic Annotations* – Gene lookup by Gene symbol or HUGO symbol and Gene Ontology Lookup by Protein.
  - *Pathways* – Pathway search by Pathway name, Pathway by Source/Provider, Pathway by Protein, or Gene symbol
  - *Cancer Gene Index* – Search for Diseases or Agents associated with a Gene
- caBIO iPhone App – The caBIO iPhone App allows researchers to query for genomic annotations (genes, chromosomes, pathways, diseases, agents, etc.) from the Apple iPhone. For example, leveraging the caBIO iPhone App, a researcher can query for information on BRCA1, a gene associated with breast and ovarian cancers, and retrieve information on the gene and its location on the chromosome, a list and visualization of cellular pathways in which the gene participates in, a list of diseases associated with BRCA1 with supporting evidence information, and agents (drugs and other therapies) used to target the gene. Having access to this information via a mobile device enables on-the-spot hypothesis testing and fosters scientific collaboration through "tele" research.

caBIO for Technical Users

caBIO Technical User Scenarios

caBIO provides the IT developer community with access to a variety of integrated biomedical data to address molecular annotation needs. Below are example high-level user scenarios demonstrating potential uses of caBIO.

A bioinformatics engineer is developing an application that allows researchers to find all of the diseases associated with a specific gene and obtain a list of agents (drugs) that target the gene. The bioinformatics engineer uses caBIO to retrieve a list of diseases and agents associated with a gene.

A translational research system provides access to the results of a microarray experiment. The system needs to be extended to display annotations for each microarray reporter used in an experiment. The development team extends the system by leveraging caBIO to obtain a list of genes for each microarray reporter by microarray and reporter name.

A gene expression analysis tool allows users to analyze gene expression data. The tool needs to be extended to display annotations associated with a specific gene in a cellular pathway. The development team extends the tool by leveraging caBIO to retrieve detailed marker annotations for a given gene symbol.

caBIO Programmatic Access

caBIO data can be accessed leveraging a variety of Application Programming Interfaces (APIs) including a Java API, Web Services API, RESTful Services, and caBIG® Grid Services. caBIO also provides a number of utilities to search for and manipulate data programmatically.

Java API

Java-based clients communicate with caBIO via the Java API which contains the domain objects provided by the caBIO cabioXX-beans.jar file. Developers can download the caBIO client API and directly use its packages in their programs. Each Java bean object class in the caBIO domain packages acts as its own search criteria object and provides methods for the retrieval of attributes and related domain objects.

Web Services API

Non-Java based applications can communicate with caBIO via web services using the SOAP protocol. Currently, caBIO has a Python API which makes use of the web service.

RESTful Services

Also called the REST (Representational State Transfer) API, this interface allows users to query over HTTP and receive objects as XML.

caBIO Home provides a convenient way to search and retrieve data via HTTP queries. By default, it returns HTML objects, but the generated

query can be edited to return the XML object by simply changing the "...//GetHTML?query=..." to "...//GetXML?query=..." in caBIO Home Page results.

All caBIO objects can be transformed into XML, and XSL/XSLT is used to present data in documents, web pages or other interfaces.

See the page REST API with Ajax for information on using the REST API for quick access to caBIO data from your own web application.

caBIO Grid Service

caBIO data can also be obtained via caBIO Grid Services on caGrid. The current caBIO grid services are found on the caGrid Portal.

caBIO Utilities APIs

caBIO provides the following utility APIs.

- *XML Utility* - caBIO includes a utility class that provides the capability of converting caBIO domain objects between native Java objects and XML serializations. This allows applications using the caBIO API to quickly persist any caBIO domain object as an XML representation to share or store.
- *Manipulating Scalable Vector Graphics (SVG) Diagrams* - caBIO provides a utility class called SVGManipulator for manipulating SVG diagrams of pathways. BioCarta and its Proteomic Pathway Project (P3) provide detailed graphical renderings of pathway information. NCI's CMAP website captures pathway information from BioCarta, and transforms the downloaded image data into SVG representations that support interactive manipulation of the online images. An SVG Viewer enables the display of SVG diagrams.
- *Utilities for Text Search Manipulation* - The caBIO API provides a full text search (Google™-like search) capability known as the FreestyleLM (Freestyle Lexical Mine). The FreestyleLM search component consists of interfaces and APIs for conducting horizontal searches across caBIO domain objects. For example, when searching for "brca1", the FreestyleLM search capability will return any object that makes mention of the "brca1" gene in its attributes, such as Genes, Proteins, and Pathways.
- *Genome Range Queries* - The Range Query API provides an alternative way to search for genomic features on a chromosome, based on the physical locations. The API is capable of performing arbitrary range searches on any given chromosome (absolute search), or range searches around a feature of interest (relative search).
- *Array Annotation API* - This client-side convenience API is intended for users of the Microarray annotations in caBIO. The API simplifies access to bulk microarray annotations such as reporters, genes, and SNPs.

Downloading caBIO

The caBIO download provides open source licensing information and access to the caBIO source code, web services, Java client.jar files, programming interfaces, use cases and documentation. Please contact NCICB Application Support to obtain the latest caBIO schema, database distribution and instructions for installing the database.

Collaboration Opportunities

caBIO enables collaboration opportunities through an open development initiative that allows for the contribution of caBIO expanded capabilities. The following caBIO projects are made available through collaborative development efforts:

- caCORE Perl API - Perl API for caBIO services
- pyCaBIO - Python API for caBIO services
- AJAX with caBIO REST API - Example use of AJAX applied to the caBIO REST API
- caBIO .NET API - .NET Project site

If you are interested in participating in collaboration activities, please contact the caBIO Users group.

caBIO FAQ

If you have any general or technical questions, you can also check out caBIO FAQ web page for possible answers.

Using the caBIO Wiki

This is the wiki home page for the *caBIO* product. caBIO developers may edit this page and any user may add a comment.